

# Optimization of the Lennard-Jones Parameters for a Combined *Ab Initio* Quantum Mechanical and Molecular Mechanical Potential Using the 3-21G Basis Set

MAREK FREINDORF<sup>†</sup> and JIALI GAO\*

Department of Chemistry, State University of New York at Buffalo, Buffalo, New York 14260

Received 10 January 1995; accepted 14 April 1995

## ABSTRACT

A combined *ab initio* quantum mechanical and molecular mechanical (AI-QM/MM) potential for use in molecular modeling and simulation has been described. In this article, we summarize a procedure for deriving the empirical parameters embedded in a combined QM/MM model and suggest a set of Lennard-Jones parameters for the combined *ab initio* 3-21G and MM OPLS-TIP3P (AI-3/MM) potential. Interaction energies and geometrical parameters predicted with the AI-3/MM model for over 80 hydrogen-bonded complexes of organic compounds with water were found to be in good accord with *ab initio* 6-31G(d) results. We anticipate that the AI-3/MM potential should be reasonable for use in condensed phase simulations. © 1996 by John Wiley & Sons, Inc.

## Introduction

The use of a combined quantum mechanical and molecular mechanical (QM/MM) potential in condensed phase simulations has recently become popular because of its generality and simplicity.<sup>1-7</sup> Furthermore, the combined QM/MM

method allows convenient treatment of chemical reactions in solution with inclusion of solvent polarization effects in these calculations.<sup>1,4,8-10</sup> In the combined QM/MM approach, a condensed phase system is divided into a quantum mechanical region and a molecular mechanical region. To bridge interactions between the QM and MM particles, a QM/MM interaction Hamiltonian  $\hat{H}_{qm/mm}$  is introduced, which consists of an electrostatic component  $\hat{H}_{qm/mm}^{el}$  and a van der Waals part  $\hat{H}_{qm/mm}^{vdW}$ . The latter term,  $\hat{H}_{qm/mm}^{vdW}$ , contains the only "adjustable" parameters that may be optimized to

\* Author to whom all correspondence should be addressed.

<sup>†</sup> On leave from Department of Theoretical Chemistry, Jagiellonian University, Kraków, Karasia 3, Poland.

yield the best results for intermolecular interactions in comparison with experiments or high-level *ab initio* data.<sup>11</sup> It has been shown that the combination of quantum mechanical methods and molecular mechanical force fields can provide accurate description of intermolecular interactions needed for liquid simulations and can yield excellent results for a variety of properties associated with solvation.<sup>1,8</sup>

In the past, semiempirical quantum mechanical methods have been primarily used in condensed phase simulations because these calculations are very fast and enable fluid simulations to be carried out.<sup>1-4</sup> The combined semiempirical QM/MM potentials will continue to be used widely in the future for studying large molecules of chemical and biological interest in solution. However, semiempirical QM methods, including the empirical valence bond approach, are restricted by the parameterization nature, which permits the methods to apply only to compounds that have been represented in the initial parameterization set. On the other hand, the accuracy of *ab initio* molecular orbital (MO) calculations can easily be systematically improved by using larger basis sets or correlation methods.<sup>12</sup> Computations based on density functional theory also belong to the latter category.<sup>13</sup> Thus, it is desirable to use *ab initio* methods to construct a combined QM/MM potential to improve its applicability in liquid simulations.

In this article we report a set of optimized Lennard-Jones parameters for a combined *ab initio* MO QM/MM potential employing the split valence 3-21G basis set.<sup>14</sup> The choice for the basis set in the present study has taken into account both the need for computational accuracy and computational efficiency. It has been well documented that computations using the 3-21G basis set can yield good results for molecular geometries, energetics, conformational analyses, and reaction paths.<sup>12</sup> While results obtained from *ab initio* 3-21G calculations can certainly be improved with a larger basis set, including polarization functions such as 6-31G(*d*), the tremendous increase in computer time, disk space, and memory would prevent any practical applications of such a combined QM/MM potential for fluid simulations. We note that the procedure outlined herein can be used to derive the parameters for other basis sets, as well as for post-Hartree-Fock calculations. We will continue our efforts in this area and report our results in future publications.

In this article the empirical parameters embedded in the combined *ab initio* MO 3-21G/MM potential (hereafter referred to as the AI-3/MM potential) is first optimized by reproducing *ab initio* 6-31G(*d*) hydrogen bonding distances and geometries for a set of representative bimolecular complexes. The performance of the combined AI-3/MM potential is verified through comparison with *ab initio* 6-31G(*d*) results for bimolecular complexes from a much larger number of organic compounds, representing amino acid sidechains and nucleotide bases. The theoretical background and computational details are first given, followed by results and discussion.

## Theoretical Background

Details of the theoretical background for the combined QM/MM method have been described previously.<sup>1-5</sup> Interested readers may find refs. 1 through 8 interesting, which also contain a thorough list of the early literature. However, for completeness, definitions closely related to this work will be briefly summarized here. In the present study, the organic species are represented by a closed-shell Slater determinant wave function,<sup>12</sup> while the water molecule is treated by the empirical three-site TIP3P model.<sup>15</sup> For other MM molecules, Jorgensen's optimized potentials for liquid simulations (OPLS) functions are used.<sup>16</sup> The effective Hamiltonian for this hybrid system is given by eq. (1):

$$\hat{H}_{\text{eff}} = \hat{H}_{\text{qm}}^0 + \hat{H}_{\text{qm/mm}} \quad (1)$$

where  $\hat{H}_{\text{qm/mm}}^0$  is the Hamiltonian for an isolated QM molecule in the gas phase, and  $\hat{H}_{\text{qm/mm}}$  is the QM/MM interaction Hamiltonian describing solute-solvent interactions. The latter term can be further divided into an electrostatic and a van der Waals term:

$$\begin{aligned} \hat{H}_{\text{qm/mm}} &= \hat{H}_{\text{qm/mm}}^{\text{el}} + \hat{H}_{\text{qm/mm}}^{\text{vdW}} \\ &= \left( - \sum_{s=1}^S \sum_{i=1}^{2N} \frac{q_s}{r_{si}} + \sum_{s=1}^S \sum_{m=1}^M \frac{q_s Z_m}{R_{sm}} \right) \\ &\quad + \sum_{s=1}^S \sum_{m=1}^M 4\epsilon_{sm} \left[ \left( \frac{\sigma_{sm}}{R_{sm}} \right)^{12} - \left( \frac{\sigma_{sm}}{R_{sm}} \right)^6 \right] \end{aligned} \quad (2)$$

where *i*, *m*, and *s* are, respectively, indexes for electrons, QM nuclei, and MM interaction sites; *M*

and  $S$  are the total number of QM and MM atoms;  $q_s$  is the molecular mechanics partial charge on the MM atom;  $Z_m$  is the QM nuclear charge; and  $r_{si}$  and  $R_{sm}$  are distances of electrons and nuclei from the MM sites. Standard combining rules are used to determine the parameters in the Lennard-Jones term, such that  $\epsilon_{sm} = (\epsilon_s \epsilon_m)^{1/2}$  and  $\sigma_{sm} = (\sigma_s \sigma_m)^{1/2}$ , where  $\epsilon_s$  and  $\sigma_s$  are parameters for the MM interaction sites. The  $\epsilon_m$  and  $\sigma_m$  are new empirical parameters introduced in the combined QM/MM potential, which typically depend on the atomic number and in some cases also on the hybridization state of that atom.<sup>11</sup> They are expected to be fully transferrable from one compound to another.<sup>1,8,11</sup> The requirement for the Lennard-Jones term in eq. (2) is to provide an estimate of dispersion energies between QM and MM atoms and to prevent distance collapsing between the QM electrons and MM sites.<sup>3</sup> For use of eq. (1) in liquid simulations, an additional molecular mechanical energy term will be added to describe interactions among solvent molecules.<sup>1-4,17</sup>

The introduction of the environment effects, eq. (2), into quantum mechanical calculations involves addition of the one-electron integral term in the standard Fock matrix in Hartree-Fock calculations:

$$F_{\mu\nu} = F_{\mu\nu}^0 + I_{\mu\nu} \quad (3)$$

where  $F_{\mu\nu}^0$  is the Fock matrix element for an isolated QM molecule and  $I_{\mu\nu}$  is a one-electron integral defined by eq. (4).

$$I_{\mu\nu} = - \sum_s q_s \int \frac{\Phi_\mu \Phi_\nu}{r_{sl}} d\tau_1 \quad (4)$$

Here the summation is over all solvent charge sites, and  $\Phi_\mu$  and  $\Phi_\nu$  are basis functions. Equation (4) was evaluated analytically as implemented in the GAMESS program.<sup>18</sup>

The Coulombic term between QM and MM nuclear charges in eq. (2) is computed classically. Thus, the total energy of the hybrid QM/MM system can be expressed as follows:

$$E_{\text{tot}} = \frac{1}{2} \sum_{\mu\nu} \mathbf{P}_{\mu\nu} (\mathbf{H}_{\mu\nu} + \mathbf{F}_{\mu\nu}) + E_{\text{qm}}^{\text{nuc}} + E_{\text{qm/mm}}^{\text{nuc}} + E_{\text{qm/mm}}^{\text{vdW}} \quad (5)$$

where  $\mathbf{H}$  and  $\mathbf{F}$  are the one-electron and Fock matrices, respectively, both including contribu-

tions of eq. (4);  $\mathbf{P}$  is the Hartree-Fock one-electron density matrix;  $E^{\text{nuc}}$  is the nuclear Coulombic energy for the QM and QM/MM terms; and  $E_{\text{qm/mm}}^{\text{vdW}}$  is the empirical energy of the Lennard-Jones term in eq. (2).

The hydrogen bonding interaction energies are evaluated according to eq. (6):

$$E_{\text{HB}} = E_{\text{tot}} - E_{\text{qm}}^0 \quad (6)$$

where  $E_{\text{qm}}^0$  is the energy of the isolated QM molecule in the gas phase.

## Computational Details

Combined QM/MM calculations have been performed for a series of bimolecular hydrogen-bonded complexes of organic compounds with water using a program developed in our laboratory, in which the QM energies are evaluated using the GAMESS program by Schmidt et al.<sup>18</sup> The implementation of the combined QM/MM potential into the GAMESS program was straightforward since the facilities for evaluation of the one-electron integral  $I_{\mu\nu}$  are all available. In fact, features allowing the incorporation of classical point charges have already been implemented into the latest version of the GAMESS program as part of a project to combine QM calculations with the MM CHARMM program.<sup>19</sup>

All partial geometry optimizations were executed with fixed monomer geometries at the 3-21G level for the QM molecules and the experimental geometry for water using the AI-3/MM potential. The corresponding full *ab initio* MO calculations have been carried out previously using the 6-31G(d) basis set in the process of optimizing the combined AM1/TIP3P model in our laboratory, while some are available including those of nucleotide bases in literature.<sup>11,20</sup> In general, electron correlation and basis set superposition error corrections are important for determining the energetics of hydrogen bonding complexes. However, the 6-31G(d) basis set has been repeatedly demonstrated to perform exceptionally well for hydrogen bonding interactions, perhaps due to fortuitous error cancellations.<sup>12,20</sup> The 6-31G(d) results have been the primary source of justification in developing empirical MM force fields for simulation of organic compounds and biopolymers in solu-

tion.<sup>21,22</sup> Consequently, we chose to use these results to parameterize the AI-3/MM potential. All calculations were performed on IBM RS6000/370 computers in our laboratory.

## Results and Discussion

To begin, a set of bimolecular complexes of representative organic compounds with water was selected for optimization of the Lennard-Jones parameters in eq. (2). The final values are listed in Table I for H, C, N, O, F, Cl, Li, and Na atoms. The structural arrangements for these complexes are shown in Figure 1, while the interaction energies from the 6-31G(*d*) and AI-3/MM calculations are given in Tables II and III.

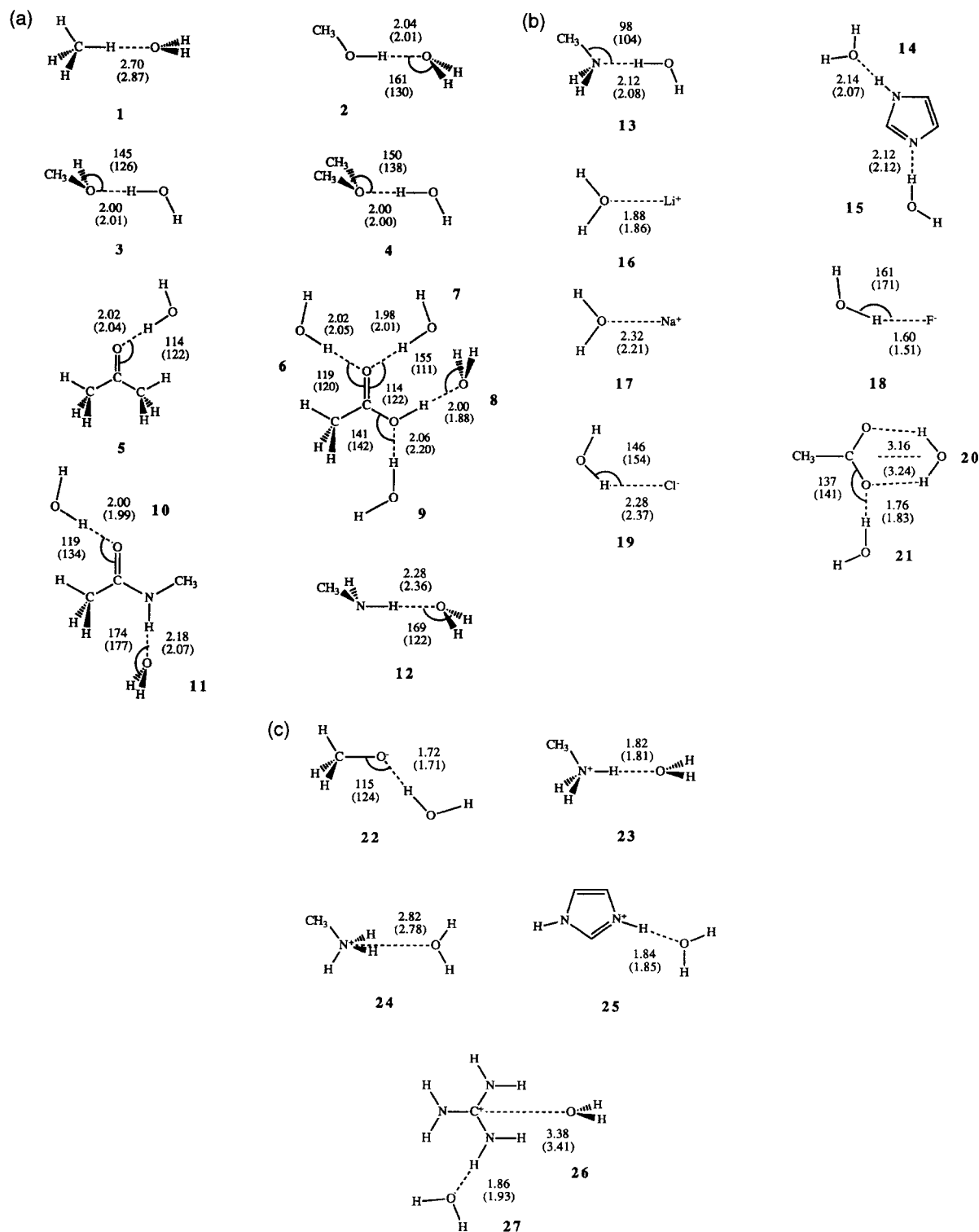
Before trends of the results are discussed, some general considerations in the parameterization process should be noted. Throughout this work, we have kept in mind and tried to generate a set of parameters that are similar to those used in empirical MM force fields, particularly those in the OPLS functions.<sup>16,20</sup> During the process, oxygen parameters were first determined based on bimolecular complexes of methanol, dimethyl ether, acetone, and acetic acid with water (2–10). In contrast to the combined semiempirical AM1/MM potential, in which different Lennard-Jones parameters for the oxygen *sp*<sup>2</sup> and *sp*<sup>3</sup> hybridizations are required,<sup>11,23</sup> we found that a single set of  $\sigma$  and  $\epsilon$  values is sufficient for all species. For the nitrogen atom, several complexes between water and

CH<sub>3</sub>NH<sub>2</sub>, imidazole, and *N*-methylacetamide (NMA) were considered (11–15). Again, a unique pair of parameters is enough to reproduce the *ab initio* results using the AI-3/MM potential. Obviously, parameters for the hydrogen atom attached to noncarbon atoms were required, which were, of course, optimized in the process. The choice of hydrocarbon parameters (aliphatic carbon and hydrogen) was somewhat ambiguous because intermolecular interactions are weak in these cases. Consequently, we decided to start with the parameters employed in the empirical MM force fields and found that only a minor adjustment of these parameters was needed to yield reasonable results for the methane–water complex (1).

Following this approach, parameters for ionic species were also obtained by considering 12 ion–water complexes (Fig. 1, 16–27); however, we found that it was necessary to use different Lennard-Jones parameters for oxygen and nitrogen atoms than those derived for the neutral molecules. On the other hand, hydrogen and carbon parameters can be described with the same values as those in the neutral complexes. The parameters for oxygen anions were particularly difficult to decide and required some special consideration. The most stable configuration in CH<sub>3</sub>CO<sub>2</sub><sup>−</sup>(H<sub>2</sub>O) complexes is the bifurcated form, 20, which has a binding energy of 20.9 kcal/mol at the 6-31G(*d*) level. Using the parameters listed in Table I, the interaction energy was predicted to be 19.0 kcal/mol. Experimentally, the enthalpy of complexation between an acetate ion and water is 15.8 kcal/mol, and its structure has been assigned to have an arrangement similar to 21.<sup>24</sup> For comparison, the *ab initio* 6-31G(*d*) and the hybrid AI-3/MM calculations yield binding energies of 16.0 and 14.0 kcal/mol, respectively. *Ab initio* 6-31G(*d*) and 6-31 + G(*d*) calculations yield binding energies of 23.7 and 21.1 kcal/mol for the CH<sub>3</sub>O(H<sub>2</sub>O)<sup>−</sup> complex 22, while the experimental value is 23.9 kcal/mol.<sup>25</sup> However, the AI-3/MM potential results in a much weaker interaction of only 17.5 kcal/mol. Nevertheless, the hydrogen bond distance predicted by the AI-3/MM potential is in accord with the *ab initio* value. Variation of the Lennard-Jones parameters has only minor effects on the interaction energy in 22. To investigate further the cause of this discrepancy, we included a set of *s* and *p* diffuse functions on nonhydrogen QM atoms in both *ab initio* and the combined AI-3/MM calculations. In addition, two enolate–water complexes

**TABLE I.**  
Optimized Lennard-Jones Parameters for a Combined *Ab Initio* 3-21G and Molecular Mechanical Potential.

| Atom               | $\sigma$ (Å) | $\epsilon$ (kcal/mol) |
|--------------------|--------------|-----------------------|
| C                  | 3.80         | 0.08                  |
| H (on C)           | 2.60         | 0.008                 |
| O                  | 3.60         | 0.15                  |
| N                  | 3.90         | 0.20                  |
| H (on heteroatoms) | 1.30         | 0.10                  |
| N <sup>+</sup>     | 3.35         | 0.15                  |
| O <sup>−</sup>     | 3.20         | 0.15                  |
| F <sup>−</sup>     | 2.80         | 0.40                  |
| Cl <sup>−</sup>    | 4.20         | 0.11                  |
| Li <sup>+</sup>    | 1.50         | 0.80                  |
| Na <sup>+</sup>    | 2.30         | 0.30                  |



**FIGURE 1.** Bimolecular complexes and geometrical parameters for complexes used in the initial parameterization of the combined *ab initio* 3-21G and molecular mechanical (Al-3 / MM) potential. In all cases, water is treated by the TIP3P model, while the other organic molecule is represented by the 3-21G basis function. Optimized Al-3 / MM values are given first, followed by full *ab initio* 6-31G(d) results in parentheses. Distances are in angstroms and angles in degrees.

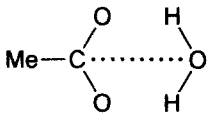
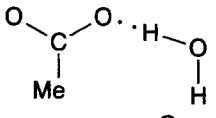
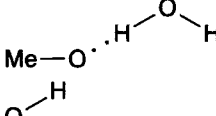
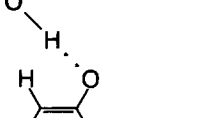
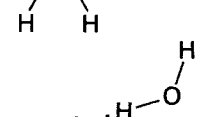
were included for comparison (Table II). This leads to a much better agreement with the corresponding *ab initio* data. Thus, use of diffuse functions seems to be important in the present combined AI-3/MM model as well as in *ab initio* MO studies of anions. For  $F^-$ ,  $Cl^-$ ,  $Li^+$ , and  $Na^+$ , only a small readjustment of the OPLS Lennard-Jones parameters was necessary. In these cases, the experimental interaction energies determined by Kebarle were used as the target in the parameter fitting.<sup>26</sup>

Overall, the agreement between the 6-31G(d) and the AI-3/MM results is excellent (Table III). The root of mean square (rms) deviation for the predicted interaction energies is 0.6 to 0.7 kcal/mol for the initial set of 27 complexes selected in the parameterization process. For structures 20–22 and the two enolate–water complexes (Table II), 6-31 + G(d) and AI-3 + /MM results are used in the comparison and the following discussion because of the critical role of diffuse functions. The accord in computed hydrogen bonding distances is also remarkable). To further justify the performance of these parameters, we have carried out partial geometry optimizations, in exactly the same manner as reported in ref. 11, for a much larger number of

bimolecular complexes that are representative of amino acid sidechains and nucleotide bases as well as other common organic functionalities. In all, a total of 83 complexes were examined (Figs. 1 and 3). The agreement between the *ab initio* 6-31G(d) binding energies and those from the present combined AI-3/MM potential is illustrated in Figure 2 with an rms deviation of only 0.5 kcal/mol.

The most significant improvement over the previous combined AM1/OPLS(TIP3P) potential as well as empirical MM force fields is a much better agreement in the optimized geometrical variables (Figs. 1 and 3).<sup>8,11,16,20</sup> In the present AI-3/MM optimization, deviations from the 6-31G(d) hydrogen bond distances are on average less than 0.1 Å for both QM donor and acceptor complexes. The only large deviation from the *ab initio* results is the  $C_6H_6-HOH$  complex (44), without which the rms deviation between the AI-3/MM and 6-31G(d) distances is only 0.07 Å (Fig. 4). In contrast, since charge polarization is much smaller in the semiempirical AM1 model,<sup>23</sup> some of the hydrogen bond distances in the combined AM1/TIP3P potential are as much as 0.4 Å shorter than the corresponding 6-31G(d) values.<sup>11</sup> Similarly, the

**TABLE II.**  
Computed Geometries and Interaction Energies for Anion – Water Complexes.<sup>a</sup>

| Complex   |        | 6-31G(d) | 6-31 + G(d) | AI-3 / MM | AI-3 + / MM <sup>b</sup> |
|---|--------|----------|-------------|-----------|--------------------------|
|  | R(C—O) | 3.20     | 3.24        | 3.14      | 3.16                     |
|   | –ΔE    | 20.9     | 18.1        | 19.0      | 19.8                     |
|   |        |          |             |           |                          |
|  | R(O—H) | 1.82     | 1.83        | 1.72      | 1.76                     |
|   | θ      | 132      | 141         | 134       | 137                      |
|   | –ΔE    | 16.0     | 14.2        | 14.0      | 13.9                     |
|  | R(O—H) | 1.70     | 1.71        | 1.68      | 1.72                     |
|   | θ      | 116      | 124         | 125       | 115                      |
|   | –ΔE    | 23.7     | 21.1        | 17.5      | 19.6                     |
|  | R(O—H) | 1.79     | 1.80        | 1.72      | 1.74                     |
|   | θ      | 126      | 137         | 146       | 127                      |
|   | –ΔE    | 17.5     | 15.3        | 14.2      | 15.7                     |
|  | R(O—H) | 1.77     | 1.78        | 1.70      | 1.72                     |
|   | θ      | 120      | 129         | 136       | 123                      |
|   | –ΔE    | 18.0     | 16.6        | 14.8      | 16.3                     |

<sup>a</sup> Units are in angstroms for distances, degrees for angles, and kcal/mol for energies.

<sup>b</sup> A set of diffuse functions is added on non-hydrogen atoms.

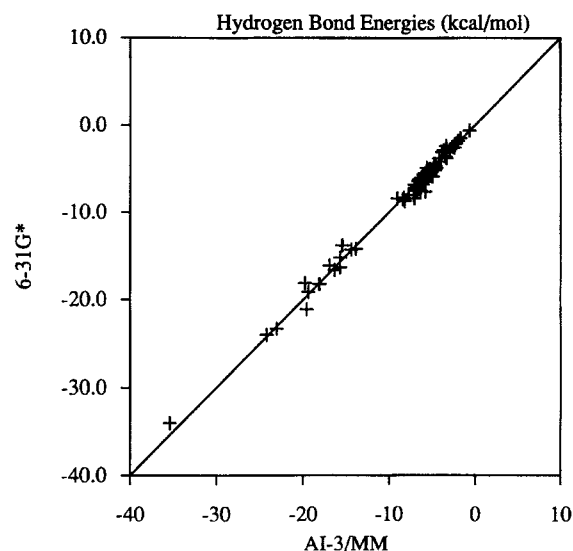
**TABLE III.**  
**Computed Hydrogen Bond Energies for Organic**  
**Compound / Water Bimolecular Complexes**  
**(kcal / mol).**

| Structure | 6-31G(d)          | AI-3 / MM | Structure | 6-31G(d)          | AI-3 / MM |
|-----------|-------------------|-----------|-----------|-------------------|-----------|
| 1         | 0.58              | 0.64      | 2         | 5.6               | 5.5       |
| 3         | 5.7               | 5.5       | 4         | 5.7               | 5.5       |
| 5         | 5.8               | 5.5       | 6         | 5.4               | 5.3       |
| 7         | 5.5               | 5.8       | 8         | 7.4 <sup>a</sup>  | 7.0       |
| 9         | 2.4               | 3.4       | 10        | 7.3               | 6.6       |
| 11        | 5.4               | 5.1       | 12        | 2.5               | 2.6       |
| 13        | 6.5               | 5.8       | 14        | 5.7               | 6.0       |
| 15        | 6.3               | 6.3       | 16        | 34 <sup>b</sup>   | 35.4      |
| 17        | 24 <sup>b</sup>   | 24.2      | 18        | 23.3 <sup>b</sup> | 23.1      |
| 19        | 14.3              | 14.4      | 20        | 18.1 <sup>a</sup> | 19.8      |
| 21        | 14.2 <sup>a</sup> | 13.9      | 22        | 21.1 <sup>a</sup> | 19.6      |
| 23        | 19.1              | 19.2      | 24        | 16.3              | 15.4      |
| 25        | 16.1              | 16.7      | 26        | 18.2              | 17.8      |
| 27        | 13.8              | 15.2      |           |                   |           |
|           |                   |           |           |                   |           |
| 28        | 5.9               | 5.8       | 29        | 5.2               | 4.7       |
| 30        | 4.9               | 5.6       | 31        | 8.5               | 8.4       |
| 32        | 6.2               | 5.8       | 33        | 7.2               | 6.7       |
| 34        | 5.9               | 4.9       | 35        | 5.5               | 5.4       |
| 36        | 6.5               | 6.0       | 37        | 6.3               | 6.5       |
| 38        | 1.5               | 1.7       | 39        | 2.6               | 2.4       |
| 40        | 5.0               | 4.8       | 41        | 5.1               | 4.6       |
| 42        | 6.4               | 6.6       | 43        | 2.1               | 2.2       |
| 44        | 3.8               | 3.2       | 45        | 4.7               | 4.6       |
| 46        | 5.6               | 5.7       | 47        | 4.4               | 4.7       |
| 48        | 1.8               | 2.0       | 49        | 3.6               | 3.4       |
| 50        | 5.4               | 5.1       | 51        | 6.8               | 7.0       |
| 52        | 6.3               | 6.6       | 53        | 8.5               | 8.3       |
| 54        | 6.7               | 6.2       | 55        | 5.6               | 5.7       |
| 56        | 8.7               | 8.1       | 57        | 7.5               | 6.7       |
| 58        | 8.4               | 9.0       | 59        | 4.9               | 4.5       |
| 60        | 5.3               | 5.4       | 61        | 5.0               | 5.2       |
| 62        | 4.3               | 4.1       | 63        | 6.0               | 6.2       |
| 64        | 5.8               | 5.6       | 65        | 7.0               | 6.7       |
| 66        | 4.5               | 4.8       | 67        | 5.9               | 5.5       |
| 68        | 4.6               | 4.7       | 69        | 3.8               | 4.2       |
| 70        | 2.8               | 3.0       | 71        | 5.2               | 5.3       |
| 72        | 5.7               | 5.4       | 73        | 2.9               | 3.7       |
| 74        | 8.3               | 8.2       | 75        | 7.2               | 7.0       |
| 76        | 8.0               | 7.1       | 77        | 3.1               | 3.9       |
| 78        | 7.6               | 5.8       | 79        | 6.8               | 6.1       |
| 80        | 8.0               | 7.7       | 81        | 7.5               | 6.3       |

<sup>a</sup> Computed with the 6-31 + G(d) basis set.

<sup>b</sup> Experimental values, refs. 24–26.

same trend is necessary in empirical MM force fields.<sup>22</sup> It is typical that hydrogen bond distances predicted using these force fields, such as the OPLS potential, are about 0.2 Å shorter than the 6-31G(d) data.<sup>16,20</sup> Shorter hydrogen bond distances in MM effective potentials are necessary to compensate for the solvent polarization effects, which are generally not treated explicitly.<sup>16,20–22</sup> Of course, in the combined QM/MM method,

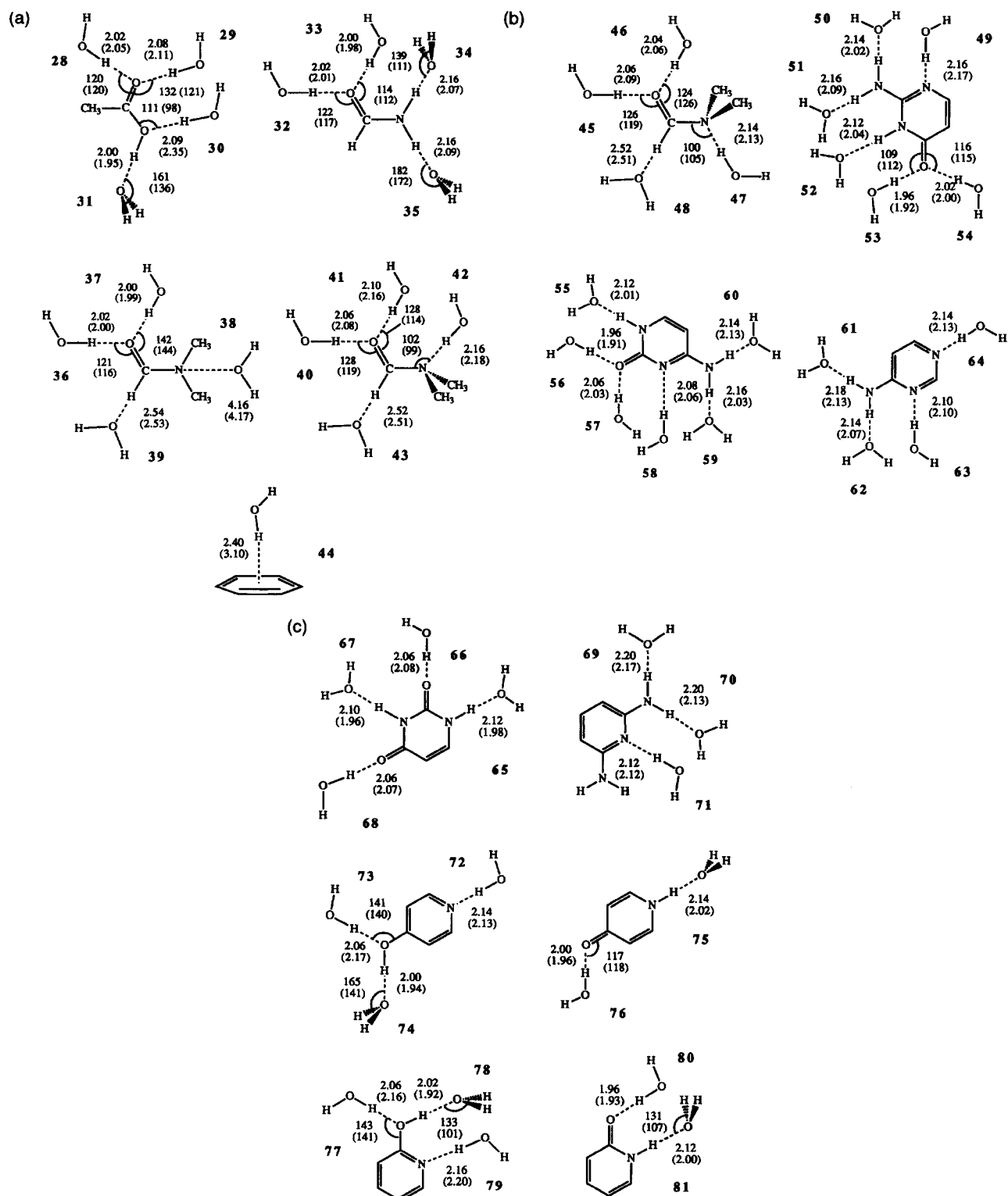


**FIGURE 2.** Comparison between *ab initio* 6-31G(d) and the AI-3 / MM interaction energies (kcal / mol).

solvent polarization effects on the QM solute molecule are naturally included in the Hartree-Fock MO calculations.<sup>1,8</sup>

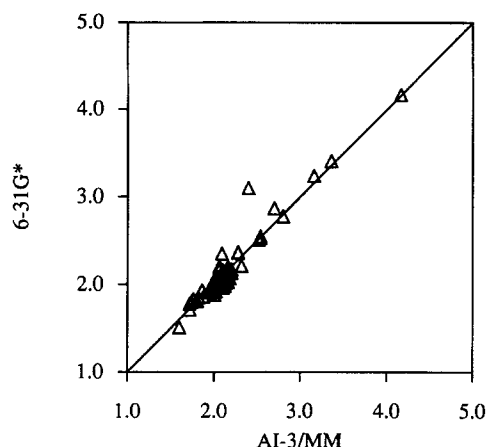
Hydrogen bond angles predicted by the AI-3/MM model also seem to show some improvement over the empirical force fields and the combined AM1/OPLS potential (Fig. 5)<sup>11</sup>; however, large deviations from the *ab initio* data still exist. In most cases, the potential energy surfaces are flat, and the energetic changes are small with variations in hydrogen bond angles.

To investigate the balance between hydrogen bond accepting and donating abilities of the QM "solute" in a QM-MM configuration, two dimer complexes are investigated (Table IV). Because donor and acceptor molecules are the same species, geometries and interaction energies should be the same whether they are treated quantum mechanically or classically; however, in the combined QM/MM model, there is an imbalance between the two arrangements. In a QM/MM dimer, one monomer is treated quantum mechanically, whereas the other monomer is treated classically. For both water and methanol dimers, QM acceptor complexes have somewhat shorter binding distances than the QM donor structures by 0.02 to 0.04 Å (Table IV). Consequently, this leads to stronger interaction energies for the QM acceptor configurations by 0.2 to 0.3 kcal/mol. The difference in these quantities can be understood by the fact that the Lennard-Jones parameters for the hydrogen atoms attached to heteroatoms are zero in

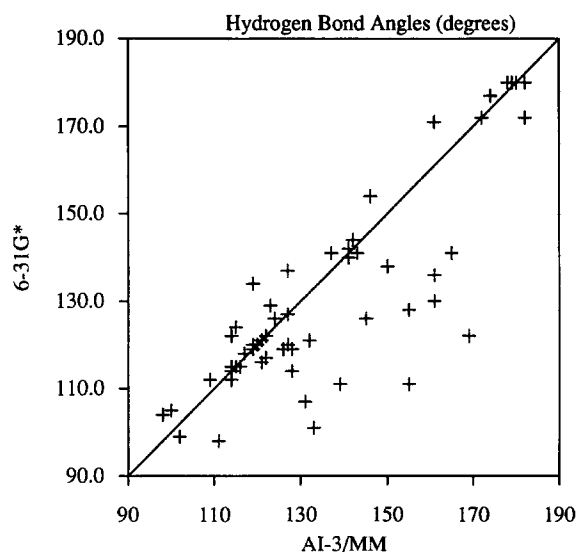


**FIGURE 3.** Structure and optimized geometrical parameters for bimolecular complexes of organic compounds with water. AI-3/MM values are given first, followed by full *ab initio* 6-31G(d) results in parentheses. Distances are in angstroms and angles in degrees.





**FIGURE 4.** Comparison between *ab initio* 6-31G(*d*) and the AI-3 / MM hydrogen bond distances (Å).



**FIGURE 5.** Comparison between *ab initio* 6-31G(*d*) and the AI-3 / MM hydrogen bond angles (degrees).

the OPLS force field, whereas the hydrogen atoms have nonzero  $\sigma$  and  $\epsilon$  values in the QM region. Thus, the van der Waals repulsion is smaller in a QM-acceptor/MM-donor hydrogen bond than in a QM-donor/MM-acceptor. Nevertheless, the small difference exhibited in the present hybrid AI-3/MM potential for the dimer complexes seems to be negligible.

The general trends of hydrogen bonding strengths predicted by the AI-3/MM potential clearly parallels those from the 6-31G(*d*) calculations (Table III). For example, among the carbonyl groups, both methods yield the same trend in increasing bonding energies: amides > acetone and acetic acid. Methanol has similar energies in complexes accepting and donating hydrogen bonds from/to a water monomer. These general features have been described previously and are obvious from the data presented in the tables and figures of this article. Interested readers are directed to ref. 11 and 22 for additional discussion.

## Conclusions

As in other computational methods, the combined quantum mechanical and molecular mechanical potential is an approximate approach that contains empirical parameters. However, the existence of these parameters provides us with an opportunity to optimize the combined QM/MM potential to obtain the best performance in describing intermolecular interactions. In this article, we presented a procedure for deriving the empirical parameters embedded in a combined QM/MM potential and suggested a set of Lennard-Jones parameters for the combined *ab initio* 3-21G and MM TIP3P-OPLS (AI-3/MM) potential. However, these parameters have not yet been tested in condensed phase simulations. The results of these tests, which are extremely demanding of computer resources, will be reported in future publications. Nevertheless, in view of the good agreement be-

**TABLE IV.** Computed Interaction Energies for Water and Methanol Dimers Using the AI-3 / MM Potential.

| Dimer  | $R(\text{O} \cdots \text{H}), \text{\AA}$ | $\theta$ (Degrees) | $-\Delta E$ (kcal / mol) |
|--|---|--------------------|--------------------------|
| HOH(qm) $\cdots$ OH <sub>2</sub> (mm)                | 2.04                                      | 158                | 5.48                     |
| H <sub>2</sub> O(qm) $\cdots$ HOH(mm)                | 2.00                                      | 143                | 5.67                     |
| CH <sub>3</sub> OH(qm) $\cdots$ OH <sub>2</sub> (mm) | 2.04                                      | 159                | 5.53                     |
| CH <sub>3</sub> (H)O(qm) $\cdots$ HOH(mm)            | 2.02                                      | 151                | 5.85                     |

tween the *ab initio* 6-31G(*d*) and the present combined AI-3/MM results for hydrogen-bonded complexes of a variety of organic compounds with water, we anticipate that the AI-3/MM potential will be a good model for describing intermolecular interactions in solution. The rms deviation from the *ab initio* 6-31G(*d*) interaction energies is only about 0.5 kcal/mol. Significantly, the predicted hydrogen bonding distances using the AI-3/MM potential are in much better agreement with the *ab initio* data (0.07 Å in rms deviation) than the combined semiempirical AM1/MM model as well as conventional MM force fields. However, it appears to be essential to include *s* and *p* diffuse functions for anionic molecules in the AI-3/MM potential.

## Acknowledgments

This work was supported in part by the National Science Foundation, the Environmental Protection Agency, and the National Institutes of Health. We thank M. W. Schmidt for providing us with a copy of the latest version of the GAMESS program. M. F. acknowledges a fellowship support through an exchange program between SUNY at Buffalo and Jagiellonian University.

## References

1. J. Gao, In *Reviews in Computational Chemistry*, K. B. Lipkowitz and D. B. Boyd, Eds., VCH, New York, Vol. 7, 1995, in press. Methods and applications of combined quantum mechanical and molecular mechanical potentials.
2. M. J. Field, In *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*, W. F. van Gunsteren, P. K. Weiner, and A. J. Wilkinson, Eds., ESCOM, Leiden, Vol. 2, 1993, p. 82.
3. M. J. Field, P. A. Bash, and M. Karplus, *J. Comp. Chem.*, **11**, 700 (1990).
4. (a) A. Warshel and M. Levitt, *J. Mol. Biol.* **103**, 227 (1976); (b) A. Warshel, *Computer Modeling of Chemical Reactions in Enzymes and Solutions*, Wiley, New York, 1991; (c) V. Luzhkov and A. Warshel, *J. Comp. Chem.*, **13**, 199 (1992).
5. R. V. Stanton, D. S. Hartsough, and K. M. Merz, Jr., *J. Phys. Chem.*, **97**, 11868 (1993).
6. D. Wei and D. R. Salahub, *Chem. Phys. Lett.* **224**, 291 (1994).
7. O. Tapia, F. Colonna, and J. G. Angyan, *J. Chim. Phys.*, **87**, 875 (1990).
8. J. Gao and X. Xia, *Science*, **258**, 631 (1992).
9. J. Gao and X. Xia, In *Structure and Reactivity in Aqueous Solution: Characterization of Chemical and Biological Systems*, C. J. Cramer and D. G. Truhlar, Eds., ACS Symp. Series 568, 1994, p. 212.
10. J. Gao, *J. Am. Chem. Soc.*, **116**, 1563 (1994).
11. J. Gao, In *Modeling the Hydrogen Bond*, D. A. Smith, Ed., ACS Symp. Series 569, 1994, p. 8.
12. W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986.
13. R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules*, Oxford University Press, London, 1989.
14. J. S. Binkley, J. A. Pople, and W. J. Hehre, *J. Am. Chem. Soc.*, **102**, 939 (1980).
15. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.*, **79**, 926 (1983).
16. W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, **110**, 1657 (1988).
17. J. Gao, *J. Phys. Chem.*, **96**, 537 (1992).
18. M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. S. Montgomery, GAMESS, version 11, 1994. See also *J. Comp. Chem.*, **14**, 1347 (1993).
19. B. R. Brooks, personal communication.
20. J. Pranata, S. G. Wierschke, and W. L. Jorgensen, *J. Am. Chem. Soc.*, **113**, 2810 (1991).
21. D. A. Pearlman and P. A. Kollman, *J. Am. Chem. Soc.*, **113**, 7167 (1991).
22. W. L. Jorgensen, *Chemtracts-Org. Chem.*, **4**, 91 (1991).
23. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, *J. Am. Chem. Soc.*, **107**, 3902 (1985).
24. M. Meot-Ner, *J. Am. Chem. Soc.*, **110**, 3854 (1988).
25. M. Meot-Ner and L. W. Sieck, *J. Am. Chem. Soc.*, **108**, 7525 (1986).
26. P. Kebarle, *Annu. Rev. Phys. Chem.*, **28**, 445 (1977).